

# Efficient Modeling and Forecasting of the Electricity Spot Price

Florian Ziel, Rick Steinert

February 28, 2014

## Abstract

The raising importance of renewable energy, especially solar and wind power, led to new impacts on the formation of electricity prices. Hence, this paper introduces an econometric model for the hourly time series of electricity prices of the EEX which incorporates specific features like renewable energy. The model consists of several sophisticated and established approaches and can be regarded as a periodic VAR-TARCH with wind power, solar power and load as influencing time series. It is able to map the distinct and well-known features of electricity prices in Germany. An efficient iteratively reweighted lasso approach is used for estimation. Moreover, it is shown that several existing models are outperformed by using the procedure developed within this paper.

## 1 Introduction

With the ongoing liberalization of electricity markets within the past decades, the volume of electricity traded via an exchange has increased profoundly. This in turn led to a raising transparency of the price for electricity. Due to companies and private households substantial dependency on the price, modeling electricity prices has become one of the great cornerstones of research in energy markets. But such modeling turns out to be at the edge of many disciplines in research. For instance, the analysis of the underlying trade mechanisms can be allocated to economics, energy production itself is a process which can be related to engineering and the rules governing the exchange of energy, especially renewable energy production, are determined by law and politics. Hence, modeling electricity prices can be a complex issue. This is also reflected in the time series, where many unusual but already stylized facts can be observed.

The proposed model within our paper tackles this complexity in several ways. Its distinctive features compared to the existent literature can be summarized to seven key facts. Our approach: 1. Models the electricity price without any data manipulation, 2. Incorporates every established stylized fact of electricity prices, 3. Provides evidence for a leverage effect within the data, 4. Proofs the effect of wind and solar power on price, 5. Accounts for specific holiday and daylight saving time effects, 6. Does not need any future information for providing accurate forecasts and 7. Uses efficient and rapid state of the art estimation techniques.

We will fit our proposed model to the hourly electricity price of the European Energy Exchange (EEX) in Leipzig for the period of 28.09.2010 up to the 25.09.2013. The data sets are obtained from the European Power Exchange (EPEX) at [www.epexspot.com](http://www.epexspot.com) for

the hourly day-ahead spot price data of Germany/Austria, from the European Network of Transmission System Operators for Electricity at [www.entsoe.eu](http://www.entsoe.eu) for the hourly load data of Germany and from the Transparency Platform of the EEX at [www.transparency.eex.com](http://www.transparency.eex.com) for the hourly wind and solar power feed-in for Germany.<sup>1</sup>

Our paper is organized as follows. In section 2 we give an overview about the setting in which electricity exchange takes place and name the distinct challenges occurring in modeling the time series. The following section sets up our model which aims to face those challenges. Section 4 presents the estimation procedures for our model. In sections 5 and 6 we fit the model to the hourly electricity price time series of the European Energy Exchange and apply a comprehensive forecast study. The last section concludes by discussing our findings.

## 2 Challenges in modeling electricity prices

Energy markets are a rapidly changing field of the economy. Liberalization of markets and the proceeding development of the energy mix significantly account for that fact. But as different countries worldwide face different preconditions, for instance politically or climatically, their energy markets tend to have a very heterogeneous structure. Hence, findings for one country may not or only gradually be used for another country or region.

In case of the German energy market, a substantial amount of the daily demand is traded via an exchange. The trading is mainly organized by the European Energy Exchange which supplies spot market and futures market trading as well as OTC services. Spot trading takes place by continuous trading and auctions. Prices for this market can be obtained either from intraday trading or from day-ahead auction prices. The latter is represented by the EPEX spot auction price for Austria and Germany. EPEX is a member of the EEX group. Since 2008, the EEX does not prohibit negative prices (Keles et al., 2012). For the day-ahead spot market there are currently 75 market participants listed.

The non-negligible amount of market participants is very likely the aftermath of liberal energy laws in Germany, especially when renewable energy is considered. The EEG and its corresponding enactments are governmental regulations which embody this idea of liberalization.<sup>2</sup> As a consequence of those the supply of energy from renewables rose significantly within the past years. Incentives, like feed-in-tariffs for renewables granted by the government, catalyzed this development. But the growing amount of power plants for renewables had direct influence on the price for electricity, (Edenhofer et al., 2013) as such energy can be produced at a cost of almost zero (Würzburg et al., 2013).

Nevertheless, also other regulations affected the market directly. For instance, transmission system operators (TSO) are obliged to sell their electricity only on the day-ahead or intraday spot market, if they decide to use an exchange.<sup>3</sup> Hence, especially data since the publication of this regulation in 2009 is of interest.

In economic theory of competitive markets the price of the electricity should equal its marginal cost. As the feed-in of renewable energy with zero marginal cost replaces every other energy source with higher marginal cost, the price of electricity should decline (Keles et al., 2013). This is known as the merit-order effect. Furthermore, the demand curve for electricity can be assumed to be inelastic (Sensfuß et al., 2008) as a certain

<sup>1</sup>We remark that the load and renewable data for Austria is not included in the data as our main focus is the German energy market.

<sup>2</sup>EEG stands for “Erneuerbare-Energien-Gesetz”

<sup>3</sup>According to article 2 AusglMechV

amount of power is needed regardless of the price. This gives rise to the implication that modeling the impact of renewables on the electricity price is of great necessity, as those energy sources will always lead to a modified marginal cost structure.

Empirical evidence for the reduction of electricity prices by the emergence of renewable energy was shown by many authors in the recent literature. For instance, [Woo et al. \(2011\)](#) use a regression analysis for the Texas electricity price market to examine the effect of wind power generation. Another multivariate regression approach was applied to the German and Austrian electricity price by [Würzburg et al. \(2013\)](#), where, among others, wind and solar power were examined and also led to a reduction in price. [Huisman et al. \(2013\)](#) gained equivalent results for the Nord Pool market by modeling energy supply and demand.

This relationship is also illustrated in Figure 1. The picture shows the price, load, solar power and wind power pattern for two weeks of October 2012. Whenever the combined effect of wind and solar power, represented by the yellow shaded graph, is high, the red line representing the price seems to decline considerably.

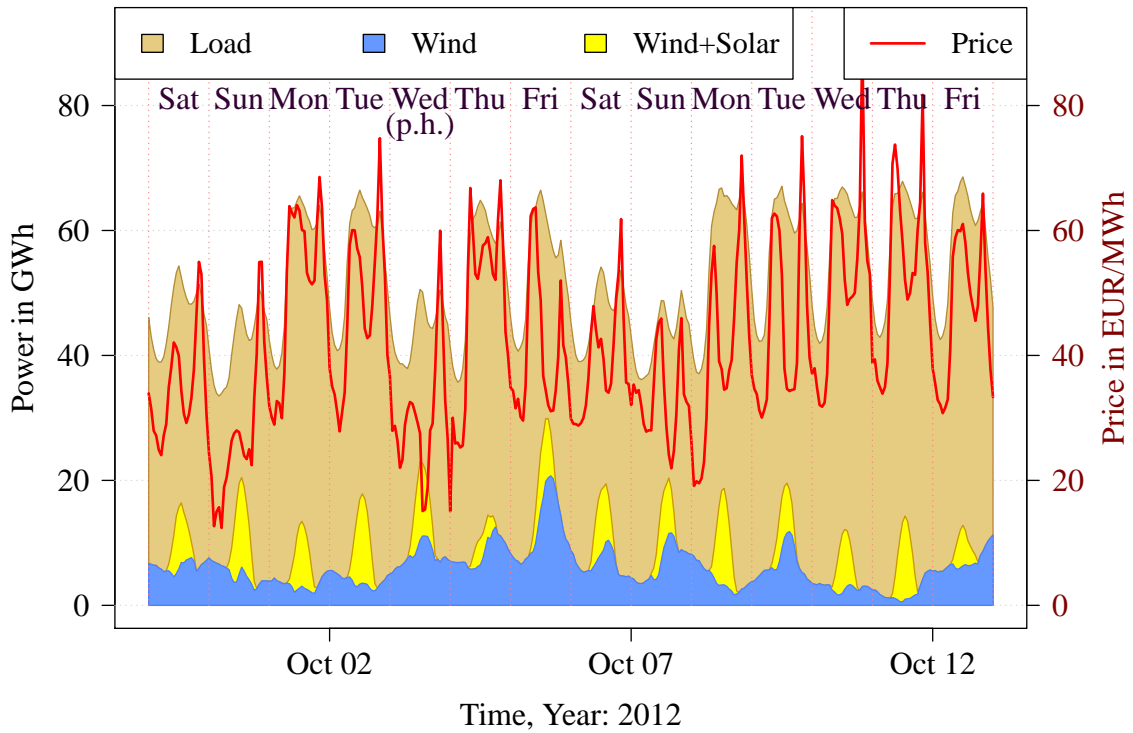


Figure 1: The hourly load, wind and solar power feed-in with the corresponding spot price. The 3rd of October is a public holiday (German unity).

But the introduction of renewables does not only lead to a price reduction effect, it can also increase the goodness-of-fit for modeling the time series of prices. Concerning this, a comprehensive study was done by [Cruz et al. \(2011\)](#). They combined sophisticated models like Holt-Winters, ARIMA and neural networks with e.g. wind power and provided evidence for their inclusion as being beneficiary for the price modeling. Also [Yan and Chowdhury \(2013\)](#), [Liebl \(2013\)](#) and [Kristiansen \(2012\)](#) yield competitive goodness-of-fit statistics by including wind power in their approaches. Using a simulation study for the dependencies of the EEX electricity price and the wind power generation [Keles et al. \(2013\)](#) were able to show the advantages of the inclusion of wind power.

The electricity price is also heavily dependent on the day of the week and their changes within the year, primarily governed by the four seasons of the year (Weron, 2006). The reason for this is twofold. First, especially the solar energy production depends by law of nature on the sunshine period which for instance is reduced in winter. Second, the daily demand for electricity is dependent on the working days, e.g. whether industrial machines are running and requiring energy or not. An example of two observed months for illustration purpose is given in Figure 1. The light-brown shaded area represents the electric load pattern of the first two October weeks in 2012. Comparing the price with the load time series indicates that both variables are positively related. As on weekends the load is usually lowered, weekends and weekdays<sup>4</sup> within the chosen period tend to exhibit different price patterns. This is also true for holidays, e.g. the German unity day on the 3rd of October, where the price and the load are both lower compared to other weekdays. A more detailed investigation of special days and phases of the day is done in section 3.

The consideration of these effects is also an important topic in the literature. The weekly dependence is usually modeled in time series analysis by incorporating the equivalent lagged value, e.g. lag 168 for hourly data. (e.g. in Kristiansen (2012)) Nevertheless, the consideration of special day or special phases of days is done only rarely. Cancelo et al. (2008) use a combined ARIMA model to forecast load and provide evidence for the inclusion of special days within the estimation. Guthrie and Videbeck (2007) calculate a periodic autoregression for the half-hourly electricity price of the New Zealand Electricity Market. They divide the day in five different time periods and show that they differ from each other.

Another appealing approach is to let the coefficients of the proposed model vary over time. For instance, Karakatsani and Bunn (2008) present a comparative analysis of models with and without time varying parameters. They proof for the British half hourly electricity price that model with time varying parameters dominate other autoregressive approaches with constant parameters. A combined model for the inclusion of holidays and time-varying coefficients was introduced by Koopman et al. (2007). They use a Reg-ARFIMA-GARCH model for the EEX, Powernext and AAX price data.

Electricity is also distinct from most other commodities as it is not efficiently storable by the means of nowadays technical equipment and adds therefore another source of risk. (Knittel and Roberts, 2005) Nevertheless, to a certain extent it can be regarded as indirectly storable because some energy sources like fuel or gas can be stored and therefore used equivalently to fulfill obligations from e.g. derivatives. (Huisman and Kilic, 2012)

The combination of almost non-storability, inelastic demand and the strong dependence on highly fluctuating energy supply cause the electricity price to be endowed with unique characteristics. Those are known as the stylized facts of electricity prices.

Most authors refer to three or even more characteristics. The most frequently used are seasonality, mean-reversion and high heteroscedastic volatility with extreme price spikes. (e.g. Weron (2006) or Eydeland and Wolyniec (2003)) For illustration purposes, the EEX electricity price of the first two months of the year 2012 are depicted in Figure 2. The pattern of the electricity price shows high price spikes in both directions, which usually last for several hours. After the impact of such a shock the price reverts to its usual daily level.

Empirical consideration of negative prices in the EEX electricity market is quite rare as they were not present until 2009. Moreover, modeling negative prices is also a difficult task, as some fitting and smoothing approaches like logarithmic transformation are not possible. Hence, some authors simply cut off or shift the time series. Nevertheless, Keles

<sup>4</sup>We use the term weekday to refer to the days from Monday to Friday

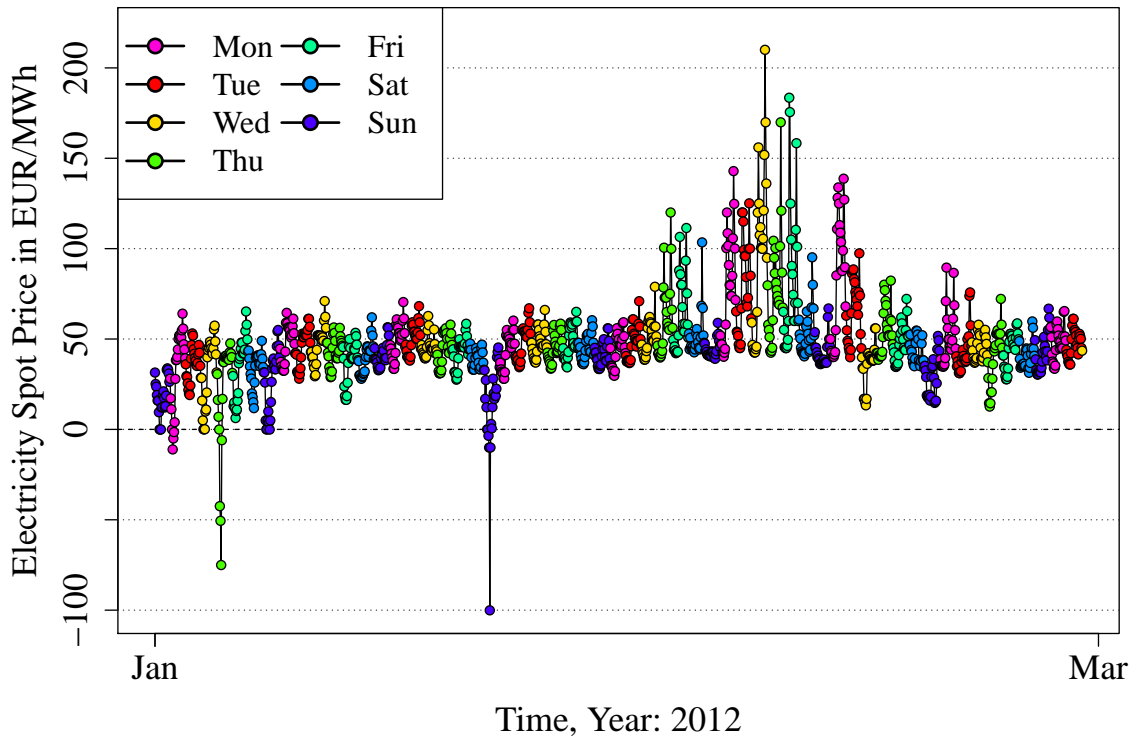


Figure 2: Hourly electricity spot price from 1st January to 1st March 2012.

et al. (2012) showed via simulation study for the EEX that the inclusion of negative prices yielded better results for their investigated models. Also Fanone et al. (2013) argue for the inclusion of negative prices.

Price spikes can be modeled in different ways. According to Christensen et al. (2012) there are mainly three different approaches common. These are specific autoregressive, Markov-regime-switching and jump diffusion models. Other authors take advantage of the combination of those models, e.g. Bordignon et al. (2013) or Escribano et al. (2011). Besides these obvious characteristics of electricity prices the recent literature argues for an appreciation of the leverage effect, which is well-known in financial economics. A leverage effect accounts for the variance of the time series reacting asymmetrically to negative and positive price shocks. (Black, 1976) The standard leverage effect has higher variance for negative than for positive price shocks. By analyzing the hourly California electricity price Knittel and Roberts (2005) were able to detect an inverse leverage effect for the time series. Within their paper they showed that the volatility responses more intense to positive than to negative shocks. This was later acknowledged for instance by Bowden and Payne (2008) and Liu and Shi (2013). Cifter (2013), however, provided evidence for the standard leverage effect by examining the daily returns of the Nord Pool electricity price market.

As a matter of fact modeling electricity prices is a complex issue. In addition to classic autoregression approaches many researchers applied methods from related scientific fields, e.g. heuristic or a combination of autoregression and other techniques. As a detailed overview for all those approaches would exhibit the scope of this paper, the interested reader is referred to the works of Weron (2006) or Aggarwal et al. (2009) who provide a comprehensive overview about the recent literature.

In the following section we will face the challenges which occur when electricity prices are examined. We introduce our own model which is tailor-made for the hourly EEX

electricity price data. It will be shown in detail, how the challenges influence our approach and how they are incorporated. We will refer to this model as periodic VAR-TARCH. It includes the hourly data of electricity load, wind and solar power feed-in. The model is capable of forecasting the price itself but also its random dependent variables.

### 3 Model for Electricity Prices

Let  $\mathbf{Y}_t$  be the considered multivariate process with dimension  $d = 3$ . So  $\mathbf{Y}_t = (Y_t^{\mathcal{P}}, Y_t^{\mathcal{L}}, Y_t^{\mathcal{R}})$  is a vector of the electricity price, the load and the renewable power feed-in at time  $t$ . Note that  $Y_t^{\mathcal{R}}$  is considered as the sum of the wind and solar power feed-in. Let  $\mathcal{I} = \{\mathcal{P}, \mathcal{L}, \mathcal{R}\}$  be the corresponding index set for the three univariate processes.

The basic model is almost equivalent to a simple vector autoregressive model. It is given by a causal solution of the following autoregressive recursion

$$Y_t^i = \mu^i(t) + \sum_{j \in \mathcal{I}} \sum_{k \in I_{i,j}} \phi_k^{i,j}(t) Y_{t-k}^j + \varepsilon_t^i \quad (1)$$

for  $i \in \mathcal{I}$ , where  $\boldsymbol{\mu}(t) = (\mu^{\mathcal{P}}(t), \mu^{\mathcal{L}}(t), \mu^{\mathcal{R}}(t))$  is a trend component,  $\boldsymbol{\varepsilon}_t = (\varepsilon_t^{\mathcal{P}}, \varepsilon_t^{\mathcal{L}}, \varepsilon_t^{\mathcal{R}})$  are noise terms with mean 0 and covariance matrix  $\Sigma_t$ , which are assumed to be independent. Note that autoregressive parameters  $\phi_k^{i,j}(t)$  may also depend on time, such as the trend  $\mu(t)$  and the matrix  $\Sigma_t$ . The considered lags are given by the index sets  $I_{i,j}$ . These index sets are crucial and determine the autoregressive dependency structure within  $\mathbf{Y}_t$ .

The considered index sets are given in Table 1. On the first view the chosen lags in  $I_{i,j}$

Index sets	contained Lags
$I_{\mathcal{P},\mathcal{P}}, I_{\mathcal{L},\mathcal{L}}$	1:49, 72:73, 96:97, 120:121, 144:145, 166:170, 192:193, 216:217, 240:241, 264:265, 288:289, 312:313, 335:337, 504:505, 672:673, 840:841
$I_{\mathcal{P},\mathcal{L}}$	1:49, 72:73, 96:97, 120:121, 144:145, 166:170, 192:193, 335:337, 504:505, 672:673
$I_{\mathcal{P},\mathcal{R}}, I_{\mathcal{L},\mathcal{R}}$	1:25, 48:49, 72:73, 96:97, 120:121, 144:145, 166:170, 192:193
$I_{\mathcal{L},\mathcal{P}}$	1, 2, 24, 25, 168, 169
$I_{\mathcal{R},\mathcal{P}}, I_{\mathcal{R},\mathcal{L}}$	-
$I_{\mathcal{R},\mathcal{R}}$	1:50, 70:74, 94:98, 118:122, 142:146, 166:170, 190:194, 214:218, 238:242, 262:266, 286:290, 310:314, 334:338

Table 1: Considered lags of the index sets  $I_{i,j}$  for  $i, j \in \mathcal{I}$ . (a:b means a, ..., b)

seem more or less random, albeit every lag is chosen with respect to economic or statistic theory.

The first lags 1, 2, ... simply model the linear dependence of the past hours. But we also added lags which concern only the dependence of previous days or weeks, such as 72 and 73. Those lags allow for a structure in the process that is similar to a multiplicative seasonal AR model. For example, the multiplicative structure of the lag polynomials in a seasonal  $\text{AR}(1) \times (2)_{24}$  is given by  $(1 - a_1 B)(1 - a_{24} B^{24} - a_{48} B^{48}) = 1 - a_1 B - a_{24} B^{24} + a_1 a_{24} B^{25} + a_{48} B^{48} + a_1 a_{48} B^{49}$  with  $B$  as backshift- resp. lag operator. Thus an AR model with autoregressive polynomial  $1 - \phi_1 B - \phi_{24} B^{24} - \phi_{25} B^{25} - \phi_{48} B^{48} - \phi_{49} B^{49}$  contains the special case of an  $\text{AR}(1) \times (2)_{24}$ , but is more general. Higher order seasonal AR processes provide the lags 72, 73 or 96, 97 and so on. Considering a weekly seasonality of 168 gives



lags like 336, 337 or 504, 505. Due to the complex dependency structure in  $\mathbf{Y}_t$  we added the mentioned lags to corresponding index sets. Note that we choose  $I_{\mathcal{L}, \mathcal{P}}$  rather sparse, as we assume that there is no strong influence of the past price to the load. Moreover, the renewable energy feed-in does not have any weekly seasonal structure, only a daily and a yearly one. Thus,  $I_{\mathcal{R}, \mathcal{R}}$  does not contain higher order weekly based lags like 504 and 505 as e.g.  $I_{\mathcal{P}, \mathcal{P}}$ . Further, we assume that the renewable energy feed-in does not depend on the past price and load. The considered choices of lags can also be underlined by considering the sample partial autocorrelation of  $\mathbf{Y}_t$ .

The trend component  $\boldsymbol{\mu}(t)$  is modeled by a linear combination of basis functions plus a linear trend, so we have

$$\mu^i(t) = \mu_0^i + \mu_{\text{lin.}}^i t + \sum_{j=1}^N \mu_j^i B_j^{\mu^i}(t) \quad (2)$$

for  $i \in \mathcal{I}$  with parameters  $\mu_j^i$  and basis functions  $B_j^{\mu^i}(t)$ . Note that in general the mean  $\mathbb{E}(\mathbf{Y}_t)$  of  $\mathbf{Y}_t$  is not equal to the mean component  $\boldsymbol{\mu}_t$ . The structure of the considered basis functions is inspired by periodic B-splines as used in (Harvey and Koopman, 1993), instead of the often used Fourier basis functions. The B-spline approach is using local basis functions that provide more flexibility, especially by modeling time series with non-stationary impacts as the electricity price or load.

We consider periodic B-splines with a degree of  $D = 3$  that provides the popular cubic splines which are twice continuously differentiable. Note that  $\sum_{j=1}^N B_j^{\mu^i}(t)$  is constant. Hence, we always drop the first component whenever we use periodic basis functions to avoid singularities as we include the corresponding intercept  $\mu_0^i$ . Moreover, we consider equidistant knots with distance  $d_K$  and a seasonality of  $S = 168$  which represents a week. To get results which are better to interpret it is useful to choose  $d_K$  as an integer valued factor of  $S$ . Thus there are  $\frac{S}{d_K}$  parameters added. For our implementations we consider  $d_K = 4$  for  $\mu^i$  and  $i \in \{\mathcal{P}, \mathcal{L}\}$ , so there would be 42 parameters added. This is quite a lot, but it is suitable to capture the mean behavior during the week. By using this proposed model for  $\boldsymbol{\mu}(t)$  we have smooth changing over the days which changes during the week as well. So for example, the values for Sunday at 8 am are different to Sunday 9 am or Monday 8 am, but are equal for every Sunday at 8 am.

Nevertheless, we also observe that the mean behavior of at least some days is very similar to other days. This finding was embedded in our model by proposing the assumption of some days having equal hourly mean components, which in turn reduces the number of parameters in the model. For example, the mean at Tuesday 8 am is assumed to be the same as on a Wednesday 8 am. Finally we consider 5 groups of parameter restrictions: Saturday, Sunday, Monday from 0 am to 12 am, Friday from 12 am to 12 pm and the core week from Monday 12 am to Friday 12 am. The Monday morning and Friday afternoon resp. evening has to be considered to get an appropriate phase-in resp. phase-out of the weekend. Figure 3 motivates our assumption. Friday, represented as the cyan colored line, leaves the typical structure of the other days excluding weekend approximately at 12 am. This is almost the same for Monday, with the difference that the Monday needs 7 to 8 hours to reach an equivalent structure compared to the other weekdays. For maintaining equidistant phase-in and phase-out periods we set both periods to 12 hours. The considered basis coefficient structure of a normal week for  $d_K = 4$  is visualized in Figure 4. Note that due to the parameter restriction there are 24 parameters in total for  $d_K = 4$  (instead of 42 parameters), but we add only 23, as the first component is dropped.

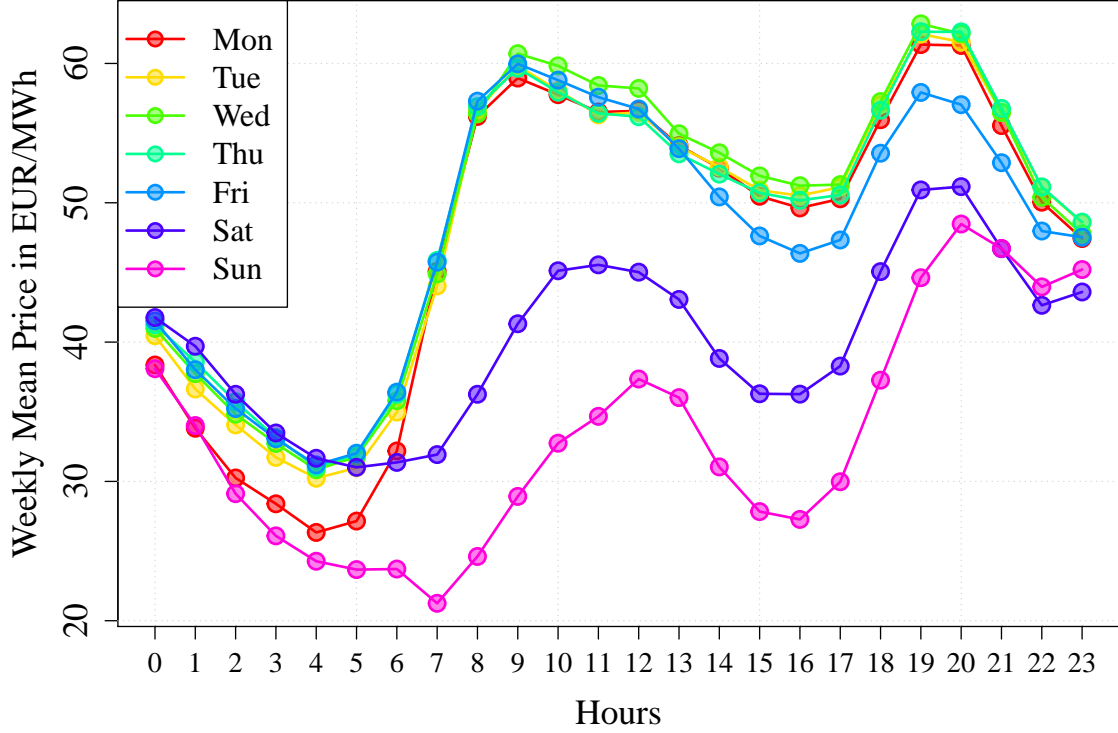


Figure 3: Average price in EUR/MWh for every hour, separated by every day of the week.

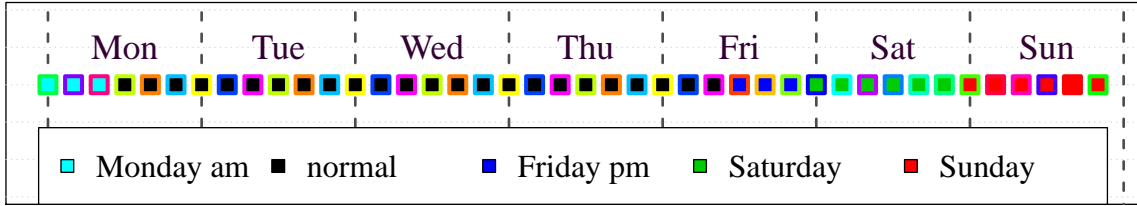


Figure 4: Structure of the created basis function in a common week for  $d_K = 4$ . Every different symbol represents one of the 24 different basis functions.

Furthermore, there are significant impacts of public holidays, as mentioned by several authors. We consider every German public holiday as a Sunday. In addition to that there are some local public holidays in Germany which are celebrated only by some areas. Every local public holiday, that concern at least 25% of the population and is not on a Sunday, is treated as a Saturday as well. Moreover, we consider Christmas Eve and New Years Eve as local public holidays. Consequently we take the 12 hours after a (local) public holiday as Monday-am and the 12 hours before every (local) public holiday as Friday-pm. Of course this applies only if the day itself is not a Sunday, Saturday or (local) public holiday. An overview of the explained day-grouping is given in Table 2.

Furthermore, we can observe an annual pattern within the data, most distinct for the load and solar data. Hence, we added B-splines with a yearly periodicity of  $24 \times 365.24 = 8765.76$ . As we assume that the meteorologic processes have a smoother changing over the year we added only 6 basis functions, whereas we take 12 basis functions for the price and load, which leads to corresponding  $d_K^{\text{annual}}$  of 1460.96 and 730.48, respectively. For the wind and solar component we assumed no weekly structure, only a daily and a yearly one. Therefore we construct basis function for the daily structure with a periodicity of 24 and annual basis functions. We choose  $d_K^{\text{daily}} = 4$ , thus there are 6 basis functions added.



Group	Contained Days	Number of parameters
<i>full off</i>	Sundays and public holidays	$24d_{\mathcal{K}}^{-1}$
<i>semi off</i>	Saturdays and local public holiday if not <i>full off</i>	$24d_{\mathcal{K}}^{-1}$
<i>phase in</i>	12 hours before <i>full</i> or <i>semi off</i> if not <i>full</i> or <i>semi off</i>	$12d_{\mathcal{K}}^{-1}$
<i>phase out</i>	12 hours after <i>full</i> or <i>semi off</i> if not <i>full</i> or <i>semi off</i>	$12d_{\mathcal{K}}^{-1}$
<i>normal</i>	others	$24d_{\mathcal{K}}^{-1}$

Table 2: Group structure of the days.

Additionally we take interactions between the yearly and daily components into account, i.e. how the daily cycle behavior changes over the year. This seems important as for instance the length of the sunshine period of the day is changing over the year and should have a strong influence. This is modeled by a multiplication of both components. As last remark to the basis functions we want to mention that the linear trend characterized by  $\mu_{\text{lin.}}$  in (2) can be regarded as a basis function as well.

Furthermore in equation (1) the parameters  $\phi_k^{i,j}(t)$  might also depend on time. So the autoregressive behavior can also change over time, especially in a seasonal manner. A related approach was also used by Koopman et al. (2007) and Bosco et al. (2007) for modeling daily electricity prices, and by Guthrie and Videbeck (2007) for intra-day prices. We assume that this dependence has a structure as in (2), so for  $\phi_k^{i,j}(t)$  it is given by

$$\phi_k^{i,j}(t) = \phi_{k,0}^{i,j} + \sum_{j=1}^n \phi_{k,j} B_j^{\phi_k^{i,j}}(t).$$

Again we have to choose the basis functions  $B_j^{\phi_k^{i,j}}(t)$ , but under a heuristic point of view it is beneficiary to take the same basis functions as for  $\mu^i$ , which is implemented in our approach. Note that from the regression point of view this corresponds to the (multiplicative) interaction between the mean  $\mu(t)$  and the corresponding lag  $\phi_k^{i,j}$ . But unfortunately this would expand the parameter space enormously, if we assume that for every possible lag  $k \in I_{i,j}$  the coefficient  $\phi_k^{i,j}(t)$  changes over time. Therefore, we assume that only the most important coefficients  $L_{i,j} \subseteq I_{i,j}$  change over time. The considered ones are summarized in Table 3, the other ones are empty.

Index sets	Lags
$L_{\mathcal{P},\mathcal{P}}, L_{\mathcal{L},\mathcal{L}}$	1, 2, 24, 25, 168, 169
$L_{\mathcal{R},\mathcal{R}}$	1, 2, 23, 24, 25

Table 3: Considered lags of the index sets  $L_{i,i}$  for  $i \in \mathcal{I}$ .

Furthermore, there is a special impact on  $Y_t$  that isn't modeled so far. Due to the daylight saving time we have a time shift in March and October. As we removed the added hour in October and extrapolated the missing hour in March, there is no problem in modeling the electricity spot price as the population adjusts very quickly their every day behavior. The same holds true for the electricity load if we assume that the influence of regenerative power feed-in is negligible. But for the wind and solar power feed-in it looks different. In winter the sun peak is at 12 am, in summer it is at 1 pm. This has an impact especially on the solar power. Figure 5 illustrates the this effect on the observed solar power feed-in in the morning. It can be observed that right after the time shift in

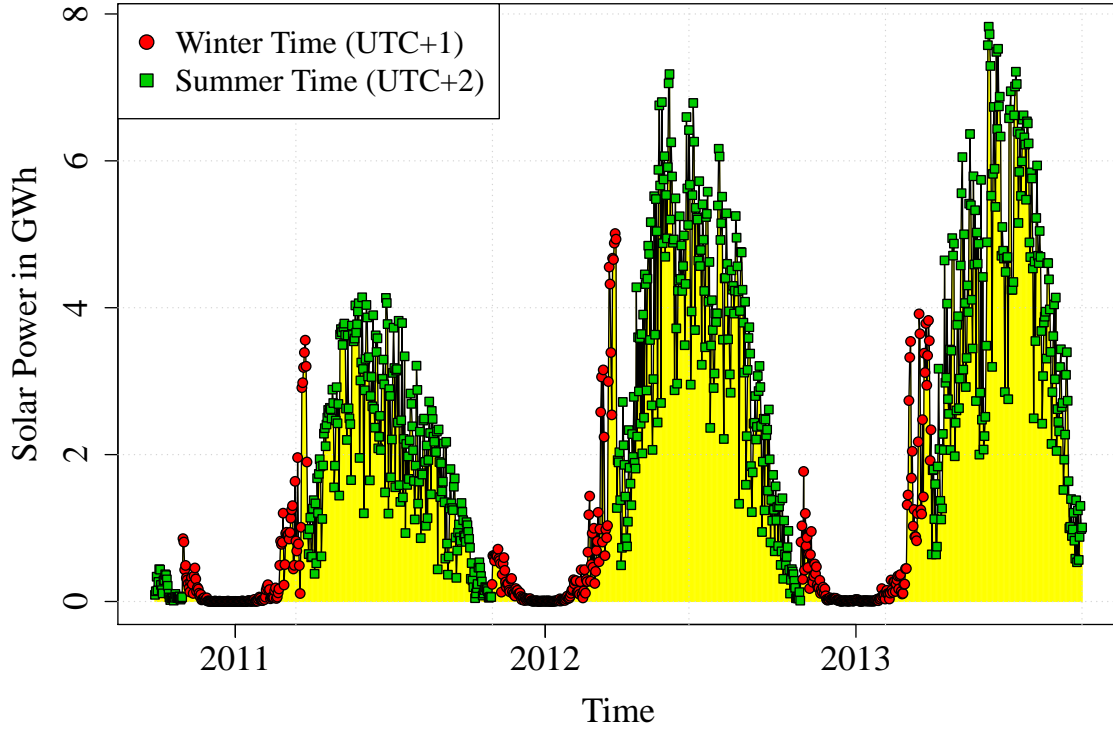


Figure 5: Solar Power feed-in at 8:00 am.

March there is an downwards jump in the solar power feed-in present and after the time shift in October a jump upwards occurs. We will model this behavior by a shift in the considered basis functions. Therefore, let  $\mathbb{T}$  denote the set of all time points with summer time. Then the considered time adjusted basis functions are given by

$$B_j^{i,\mathbb{T}}(t) = \mathbf{1}_{\{t \notin \mathbb{T}\}} B_j^i(t) + \mathbf{1}_{\{t \in \mathbb{T}\}} B_j^i(t+1).$$

This shifting is applied to all basis function of the wind and solar power feed-in.

But also the variance structure of the model is of relevance. Here we assume that  $\Sigma_t = \text{diag}(\boldsymbol{\sigma}_t) = \text{diag}(\sigma_t^P, \sigma_t^L, \sigma_t^R)$ . The error term  $\boldsymbol{\varepsilon}_t$  can be written as

$$\varepsilon_t^i = \sigma_t^i Z_t^i \text{ where } (Z_t^i)_{t \in \mathbb{Z}} \text{ is i.i.d. with } \mathbb{E}(Z_t^i) = 0 \text{ and } \text{Var}(Z_t^i) = 1.$$

As modeled in (Keles et al., 2012) we want to model conditional heteroscedasticity by an ARCH resp. GARCH type model. But as the price process  $\mathbf{Y}_t$  is heavy tailed, the residual process  $\boldsymbol{\varepsilon}_t$  is likely heavy tailed with a small tail index, too. We noticed that it is not suitable to consider square of  $\varepsilon_t^i$  as done in a typical GARCH process, because it gets even more heavy tailed. This halves the corresponding tail index and the convergence of the estimators will get worse. Instead of taking squares we use the absolute value  $|\varepsilon_t^i|$  which keeps the tail index on the same level. Some GARCH-type processes already use this structure, e.g. the TARCH model by Rabemananjara and Zakoian (1993). Additionally, a TARCH resp. TGARCH model can handle the so called leverage effect, so that negative and positive past residuals have different influence on the volatility. We assume a TARCH process for  $\sigma_t$  so that

$$\sigma_t^i = \alpha_0^i(t) + \sum_{k \in J_i} \alpha_k^{+,i} \varepsilon_{t-k}^{+,i} + \alpha_k^{-,i} \varepsilon_{t-k}^{-,i}$$

holds with  $\alpha_k^{+,i}, \alpha_k^{-,i} \geq 0$  where  $\varepsilon_k^{+,i} = \max\{\varepsilon_k^i, 0\}$  and  $\varepsilon_{t-k}^{-,i} = \max\{-\varepsilon_{t-k}^i, 0\}$ . For the TARCH process we will assume that only the trend component  $\alpha_0^i(t)$  varies seasonally

over time. We consider the same basis functions as for the corresponding  $\mu^i(t)$ , but without the linear trend, as it could conflict the positivity constraint to the parameters. Note that the estimation of  $\alpha_k^{+,i}$  and  $\alpha_k^{-,i}$  can be done by using the recursion

$$|\varepsilon_t^i| = \tilde{\alpha}_0^i(t) + \sum_{k \in J_i} \tilde{\alpha}_k^{+,i} \varepsilon_{t-k}^{+,i} + \tilde{\alpha}_k^{-,i} \varepsilon_{t-k}^{-,i} + v_t^i \quad (3)$$

where  $\tilde{\alpha}_0^i(t) = \gamma^i \alpha_0^i(t)$ ,  $\tilde{\alpha}_k^{+,i} = \gamma^i \alpha_k^{+,i}$ ,  $\tilde{\alpha}_k^{-,i} = \gamma^i \alpha_k^{-,i}$  and  $v_t^i = \sigma_t^i(|Z_t^i| - \gamma^i)$  with  $\gamma^i = \mathbb{E}|Z_t^i|$ . Here  $v_t^i$  is a weak white noise process with  $\mathbb{E}(v_t^i) = 0$ . The fitted values  $\tilde{\sigma}_t^i$  of equation (3) are proportional to the  $\sigma_t^i$  up to the constant  $\gamma^i$ .  $\gamma^i$  is the first absolute moment of  $\varepsilon_t^i$  which might help to characterize the distribution of  $\varepsilon_t^i$ . If  $\varepsilon_t^i$  follows a normal distribution it is exactly  $\sqrt{2\pi^{-1}} \approx 0.798$ , whereas e.g. the standardized t-distributions have a larger first absolute moment.

The index sets  $J_i$  that we use are given in Table 4, they contain the typical discussed lags.

Index sets	Lags
$J_{\mathcal{P}}, J_{\mathcal{L}}$	1:49, 71:73, 95:97, 119:121, 143:145, 167:169, 191:193, 335:337
$J_{\mathcal{R}}$	1:49, 71:73, 95:97, 119:121, 143:145, 167:169, 191:193

Table 4: Considered lags of the index sets  $J_i$  with  $i \in \mathcal{I}$  for the TARCH part.

Finally we want to sum up that the model contains approximately 1500 parameters, that are given in Table (5). There are about 1000 parameters in the 3-dimensional mean model and 500 parameters in the TARCH model for the volatility.

number of parameters	$\mu^i(t)$	$I_{i,\mathcal{P}}$	$I_{i,\mathcal{L}}$	$I_{i,\mathcal{R}}$	seasonal coefficients	$\phi^{i,j}$	$\sigma_t^i$	sum ( $p_i$ )
$i = \mathcal{P}$	36	85	71	64	$6 \times 35$		175	621
$i = \mathcal{L}$	36	6	85	64	$6 \times 35$		175	576
$i = \mathcal{R}$	37	-	-	110	$5 \times 36$		170	497

Table 5: Summary table of the amount of used parameters  $p_i$  of the model given in (1) for  $i \in \mathcal{I}$ .

## 4 Estimation Method

For the estimation of the parameters we will consider a variation of the iteratively weighted least squares approach as used for example in Mbamalu and El-Hawary (1993) or Mak et al. (1997). So basically we apply a weighted least squares estimation of model (1). Then we use the corresponding estimated residuals to estimate the volatility which is used to compute new weights for reestimating model (1).

But as we have a lot of regressors it is likely that some of them have no significant impact on the model, they rather increase spurious effects than improving the model. To work around this problem we suggest a lasso (least absolute shrinkage and selection operator) approach. It was introduced by Tibshirani (1996) in the context of shrinkage and selection in regression models and was recently applied by Hsu et al. (2008) and Ren and Zhang (2010) in the context of vector autoregressive models. Here we apply the least angle regression (LARS) estimation algorithm of Efron et al. (2004) for estimating the

parameters combined with the iterative reweighting computed as the testing procedure. As selection criterion we consider the in Efron et al. (2004) suggested  $C_p$  criterion, which corresponds to the Akaike information criterion (AIC) from time series analysis, resp. cross-validation in regression analysis.

For the next weight estimation step and thus the conditional volatility part of the process we have to ensure that all parameters take non-negative values. For solving this non-negative least squares problem we use the NNLS algorithm as described by Lawson and Hanson (1995). Note that this estimation technique can be seen as a parameter selection approach as well, as some parameter are likely estimated to be 0.

The general estimation scheme is given by

- 1) Set the initial  $d \times n$  dimensional weight matrix  $\mathbf{W} = (\mathbf{1}, \dots, \mathbf{1})$  and  $i = 1$ .
- 2) Estimate (1) using LARS-lasso method with weights  $\mathbf{W}$ .
- 3) Estimate (3) using  $|\hat{\varepsilon}_t|$  as absolute estimated residuals from 2) using the NNLS algorithm.
- 4) Redefine  $\mathbf{W} = (\mathbf{w}^1, \dots, \mathbf{w}^d)$  by  $\mathbf{w}^i = \left( \left( \hat{\sigma}_1^i \right)^{-2}, \dots, \left( \hat{\sigma}_n^i \right)^{-2} \right)$  with  $\hat{\sigma}_t = \left( \hat{\sigma}_t^1, \dots, \hat{\sigma}_t^d \right)$  as fitted values from 3).
- 5) if  $i < R$  then  $i = i + 1$  and back to 2) otherwise stop the algorithm.

Note that before performing the LARS algorithm on the weighted regressor matrix, we standardize their columns. Consider that in general  $\hat{\sigma}_t$  in step 4 will not estimate  $\sigma_t$ . But we have  $\sigma_t = \frac{1}{\gamma^i} \tilde{\sigma}_t^i$ . Using the plug-in estimator for  $\gamma^i$ , given by  $\hat{\gamma}^i = \frac{1}{n} \sum_{t=1}^n \frac{|\hat{\varepsilon}_t^i|}{\hat{\sigma}_t^i}$ , we can estimate  $\sigma_t$  by  $\hat{\sigma}_t^i = \frac{1}{\hat{\gamma}^i} \hat{\sigma}_t^i$ .

In the empirical application we noticed that this algorithm converges very fast, so a small number of iterations  $R$  seems to be sufficient. In our case we choose  $R = 3$ .

The used estimation technique has the huge advantage of being fast, even if hundreds of regressors are included in the model. The computational complexity of the dominating LARS algorithm is the same as for a common OLS estimation, which is  $\mathcal{O}(dp^2n)$  for  $p < n$  where  $d$  is the dimension of the process,  $p$  is the maximal number of regressors in (1) and  $n$  is the number of observations. In our computations the iteratively reweighted lasso with  $R = 3$  takes approximately one minute on a 3 GHz computer, having  $n$  about 25000,  $p$  about 450 and  $d = 3$ .

The consistency and asymptotic behaviour of the used estimator for considered process is not clear. But remember that our process behaves similar to a multi-periodic, multivariate AR-TARCH process. There is a result concerning the asymptotic normality for the least square estimate (resp. the Gaussian conditional likelihood) of the periodic, multivariate, non-linear AR( $\infty$ )-ARCH( $\infty$ ) process available, as discussed by Ziel and Schmid (2013), as a generalisation of the results from Bardet et al. (2009). Moreover the lasso-type estimators in a heteroscedastic regression setting were recently analysed in Svrien and Eric (2012) and Wagener and Dette (2013). Additionally there are some asymptotic results of the lasso estimator in an autoregressive setting. Important results are given for univariate ARX processes by Wang et al. (2007) and for the VAR model by Hsu et al. (2008).

Even though there is lack of theoretical justification to the asymptotic properties of the used estimators, we will assume asymptotic normality for them. So for the estimator

of the parameter vector  $\boldsymbol{\theta}$  of (1) we assume  $\sqrt{n}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i^*) \rightarrow N(0, \boldsymbol{\Gamma}(\boldsymbol{\theta}_i^*))$  in distribution with  $\hat{\boldsymbol{\Gamma}} = \hat{\boldsymbol{\sigma}}^2(\hat{\mathbf{G}}_i + \|\hat{\boldsymbol{\theta}}_i\|_1 \text{diag}(\hat{\boldsymbol{\theta}}_i^{-1}))^{-1} \hat{\mathbf{G}}_i (\hat{\mathbf{G}}_i + \|\hat{\boldsymbol{\theta}}_i\|_1 \text{diag}(\hat{\boldsymbol{\theta}}_i^{-1}))^{-1}$  as estimator for  $\boldsymbol{\Gamma}(\boldsymbol{\theta}_i^*)$  of the non-zero part of  $\hat{\boldsymbol{\theta}}_i$  for  $i \in \mathcal{I}$  with regression matrix  $\mathbf{X}_i$ , weight matrix  $\hat{\mathbf{W}}_i = \text{diag}((\hat{\sigma}_1^i)^{-2}, \dots, (\hat{\sigma}_n^i)^{-2})$ , weighted Gramian  $\hat{\mathbf{G}}_i = \mathbf{X}_i^\top \hat{\mathbf{W}}_i \mathbf{X}_i$ , variance estimate  $\hat{\boldsymbol{\sigma}}^2$  of the lasso and  $L_1$  norm  $\|\cdot\|_1$ .

## 5 Results

We performed the mentioned estimation technique on the full data set. For the model diagnostic the sample autocorrelation function (ACF) of the estimated standardized residuals and their absolute values are given in Figure 6.

In the left 3x3 matrix it is depicted how the standardized residuals of one time series correlate with the lagged standardized residuals of the same or any other of the three time series. The right 3x3 matrix shows the same relationship for absolute standardized residuals. For instance, the left upper ACF shows the correlation of the standardized residuals of the price with its own lagged standardized residuals. The illustration right next to it, Price & Load, shows the correlation of the standardized residuals of the price with the lagged standardized residuals of the load. Hence, we can see that the independence assumption to  $\mathbf{Z}_t$  seems to be satisfied for most of the relationships. However, the strongest serial correlation structure can be observed for the standardized residuals  $Z_t^{\mathcal{L}}$  of the load. Further, the sample autocorrelation of  $|\mathbf{Z}_t|$  is mainly zero, so that  $\mathbf{Z}_t$  seems to be quite homoscedastic.

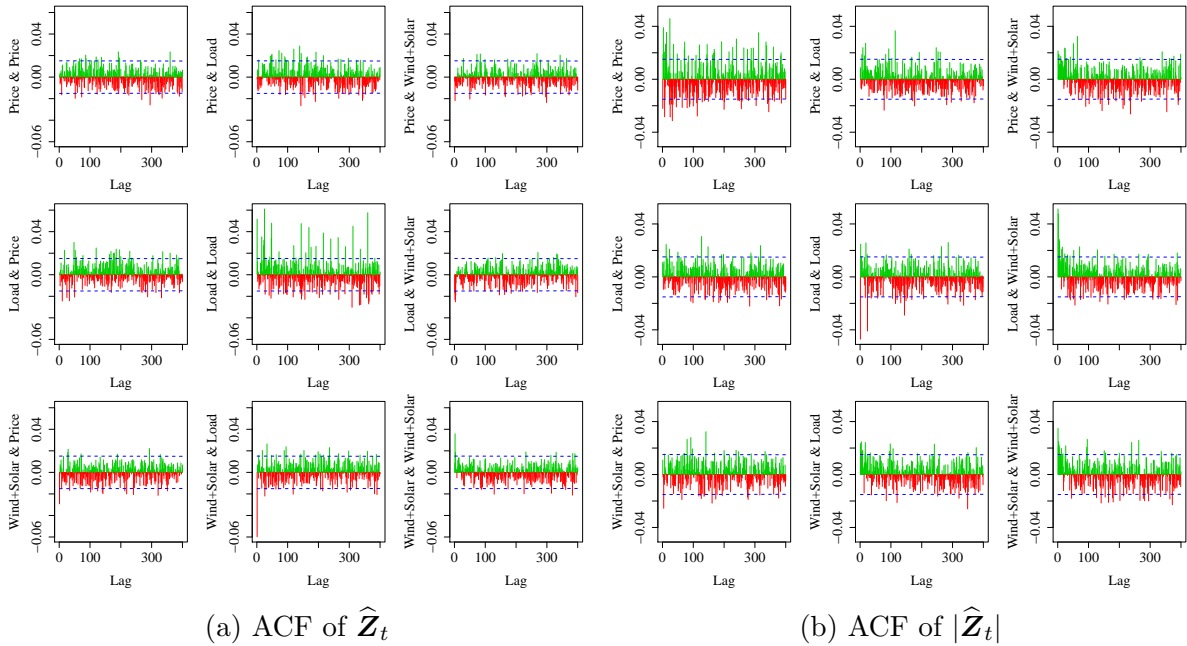


Figure 6: Sample autocorrelation function of  $\hat{\mathbf{Z}}_t$  (6a) and  $|\hat{\mathbf{Z}}_t|$  (6b)

Given the estimated model we can evaluate the t-values of every coefficient given by their estimate over its standard error which can be estimated under the asymptotic normality assumption. Roughly spoken we can say, the larger the absolute t-value of a coefficient the more significant is its impact on the process. The evaluated t-values are given in Figure 7.

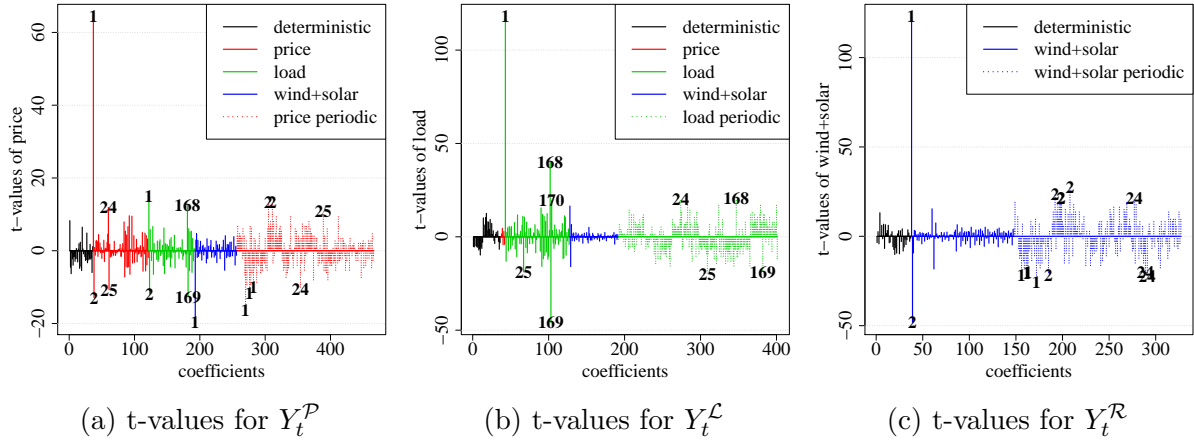


Figure 7: T-values of the estimated model for  $\mathbf{Y}_t$  distinguished by deterministic, price, load, wind+solar and periodic coefficients.

It is worth mentioning that the load and wind and solar feed-in seem to have a strong impact on the actual electricity price. The past load values (lag 1 and 2) seem to be important as well as the values one week ago (lag 168, 169). In contrast to this the impact of wind and solar on the price and the load is very important only in the short-run at lag 1. For the load we estimated a strong daily (lag 24, 25) and weekly (lag 168, 169, 170) seasonal autoregressive structure. For the solar and wind power feed-in it is interesting that the periodic structure in the first and second lag is estimated as more important than the daily lagged feed-in.

Assuming asymptotic normality allows us to compute several confidence intervals and bands. So for example for  $\mu(t)$  and time varying autoregressive parameters  $\phi_k^{i,j}(t)$ . Exemplary we illustrate the behavior of  $\mu^P(t)$  and  $\phi_1^{P,P}(t)$  in Figure 8. From this illustration it can be obtained that all coefficients have distinct weekly patterns and a seasonal structure.

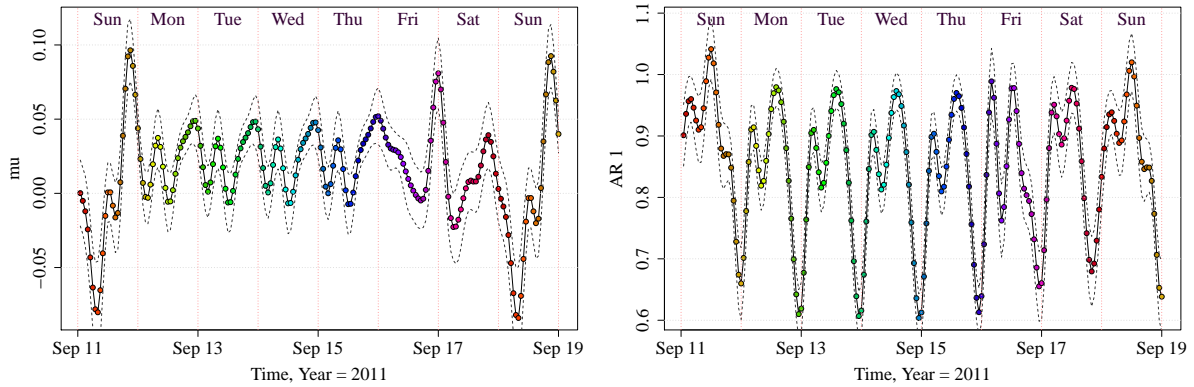


Figure 8:  $\hat{\mu}^P(t)$  and  $\hat{\phi}_1^{P,P}(t)$  within 8 days with approximate 95% confidence intervals.

Moreover, analyzing the volatility structure determined by  $\sigma_t$  reveals promising results. Using the TARCh model for  $\varepsilon_t$  allows us to decompose  $\sigma_t$  into three components. The first one is determined by the periodic coefficient  $\alpha_0(t)$ . It describes the deterministic part of the volatility and is automatically a lower bound of  $\sigma_t$ . The second and third component describe the residual impact of positive and negative residuals, respectively. Note that in an ARCH model both components are equal. The decomposition of the estimated volatilities  $\hat{\sigma}_t^i$  are given in Figure 9. For illustration purposes we selected approximately two weeks of October 2011. It can be obtained that the volatility of each of



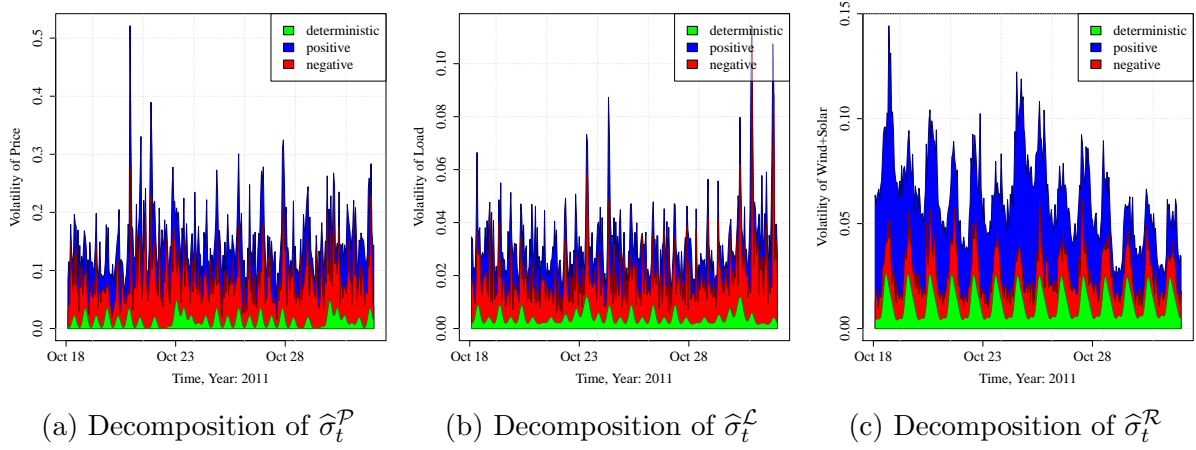


Figure 9: Decomposition of the estimated volatilities  $\hat{\sigma}_t^i$  into their deterministic seasonal part, the part influenced by positive residuals and positive ones for the price, load and wind+solar.

the three time series is influenced by a deterministic seasonal behavior. Moreover, for the price process we observe a strong leverage effect, the impact of past negative residuals is more dominant than the impact of the past positive ones. This is represented by the red shaded area in the left picture being larger than the blue shaded area, as those represent the impact of negative and positive residuals respectively. This relationship is obviously reverted for the wind and solar time series, leading to the conclusion that for this time series an inverse leverage effect is present. For the load time series both effects seem to be balanced.

Furthermore, we can perform tests to check the significance of the leverage effect. Therefore we evaluate the test  $H_0 : A_i = \sum_k \alpha_k^{+,i} - \alpha_k^{-,i} = 0$  for  $i \in \mathcal{I}$ . Table 6 provides the corresponding point estimates, standard errors, t- and p-value. As result we get a clear significant standard leverage effect for the price and a significant inverse leverage effect for the wind and solar volatility.

price				load			
$\hat{A}_{\mathcal{P}}$	$\hat{\sigma}(\hat{A}_{\mathcal{P}})$	t-val.	p-val.	$\hat{A}_{\mathcal{L}}$	$\hat{\sigma}(\hat{A}_{\mathcal{L}})$	t-val.	p-val.
-0.1899	0.0026	-3.6861	0.0002	-0.0470	0.0019	-1.0733	0.2832
wind + solar							
$\hat{A}_{\mathcal{R}}$	$\hat{\sigma}(\hat{A}_{\mathcal{R}})$	t-val.	p-val.				
0.6726	0.0026	13.0844	0.0000				

Table 6: Test leverage effect, two-tailed test.

In the same way we can perform an approximate test if an increasing load or wind-solar feed-in leads to a significantly increasing or decreasing price in the long run. This tests have the null hypothesis  $H_0 : \Phi_{\mathcal{P},\mathcal{L}} = \sum_{k \in I_{\mathcal{P},\mathcal{L}}} \phi_{k,t}^{\mathcal{P},\mathcal{L}} = 0$  resp.  $H_0 : \Phi_{\mathcal{P},\mathcal{R}} = \sum_{k \in I_{\mathcal{P},\mathcal{R}}} \phi_{k,t}^{\mathcal{P},\mathcal{R}} = 0$ . Applying this test to the data (see Table 7) it can be concluded that in the long run an increasing load leads to a significant raise in the price whereas an increasing amount of wind or solar power feed-in causes a decrease in the price. Taking the standardization of  $Y_t$  before the estimation of (1) into account we can quantify these results. So an increase of 1 GWh load raises the price by  $1.200(\pm 0.574)^5 \frac{\text{EUR}}{\text{MWh}}$  and an increase of 1 GWh wind or solar power feed-in changes the price by  $-0.745(\pm 0.538) \frac{\text{EUR}}{\text{MWh}}$ . This result is consistent

<sup>5</sup>The term in parentheses gives the symmetric 90% confidence interval

with the recent literature for quantifying the impact of renewable energy in Germany. (Würzburg et al., 2013) But note that in the past the load was decreasing, so both effects sum up and give a decreasing price in the long run. Moreover, we estimated a linear trend in  $Y_t^P$  which is negative and significantly different from zero, so over the time itself the price seems to decrease. But this effect is not that strong, over one year the price changes by  $-0.0191(\pm 0.0074)\frac{\text{EUR}}{\text{MWh}}$ . But we have to be careful with extrapolating a linear trend far in the future, as it is economically impossible for such a negative trend to last forever.

load on price				wind+solar on price			
$\hat{\Phi}_{P,L}$	$\hat{\sigma}(\hat{\Phi}_{P,L})$	t-val.	p-val.	$\hat{\Phi}_{P,R}$	$\hat{\sigma}(\hat{\Phi}_{P,R})$	t-val.	p-val.
0.0229	0.0067	3.4362	0.0006	-0.0083	0.0037	-2.2772	0.0228

Table 7: tests for effects of the regressors on the price, two-tailed test.

## 6 Forecasting

Given the estimated model (1) of the sample  $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$  we can easily do a forecast. Let  $\mathbf{Y}_t = g(\mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, \dots)$  be the representation of (1) then we can compute  $\hat{\mathbf{Y}}_{n+h}$  iteratively by

$$\hat{\mathbf{Y}}_{n+h} = g(\hat{\mathbf{Y}}_{n-1+h}, \hat{\mathbf{Y}}_{n-2+h}, \dots)$$

and defining  $\hat{\mathbf{Y}}_t := \mathbf{Y}_t$  for  $t \leq n$ .

We performed a simulation study, where we choose subsequences of the  $n = 26233$  observations. So we choose a data part  $H_l$  that is approximately two years long (exactly 17641 observations = 105 weeks + 1 hour), starting at observation  $l$ . The  $l$  is a randomly chosen number between 0 and 7920. Finally we performed a complete estimation on the data sample  $H_l$  and estimate the next  $h = 672$  hours, that is 4 weeks. This method provides a fair estimation technique, as we use the same amount of past observations for every forecast. This procedure is repeated  $N = 2000$  times, while  $\hat{\mathbf{Y}}_{h,i}$  and  $\mathbf{Y}_{h,i}$  for  $i \in \{1, \dots, N\}$  denote the corresponding predicted values and observation, respectively. Note that we are forecasting the 3 dimensional process  $\mathbf{Y}_t$ . However, in this paper we are interested in the price process  $Y_t^P$ , therefore the further discussion will be focusing on this process.

As own benchmarks we consider the weekly persistent process  $Y_t^P = Y_{t-168}^P$  as naïve model, such as two AIC selected (V)AR processes with time varying mean. The model is given by

$$\mathbf{X}_t = \boldsymbol{\mu}_i \mathbf{1}_{\{t \in (168N+i)\}} + \sum_{k=1}^p (\boldsymbol{\Phi}_k \mathbf{X}_{t-k} - \boldsymbol{\mu}_i \mathbf{1}_{\{t \in (168N+i)\}}) + \boldsymbol{\varepsilon}_t$$

where  $i \in \{1, \dots, 168\}$ . We consider the univariate choice  $\mathbf{X}_t = Y_t^P$  and the two dimensional incorporating the load as well, so  $\mathbf{X}_t = (Y_t^P, Y_t^L)$ . We also tried to include the wind and solar feed-in but the out-of-sample performance got worse. We estimated the models in a two step approach, first removing the weekly mean and second estimating a (V)AR process via Gaussian AIC selection, where the 168 parameters for the mean are ignored. For the estimation process we solve the Yule-Walker equations that provide a guaranteed stationary solution. As maximal possible order for the univariate process we choose  $p_{\max} = 1210$  and for two-dimensional one  $p_{\max} = 555$ . The estimation of the processes is very fast and done in a few seconds. However, in our empirical results the chosen order  $p$  of the AR process is chosen usually about 800, whereas the in the two-dimensional

VAR case it is usually about 400, so both cover the weekly mean and conditional mean behavior.

Moreover, we tried to use as many models from the recent literature for a benchmark as possible. But unfortunately finding the right competitor is difficult for several reasons. First of all many authors only consider positive observations or delete some chunks of the data (e.g. outliers, holidays, etc.). Second, some of them consider information from random regressors such as the load as known, which in a real world situation would not be the case. And finally there are some models, especially from machine learning, that simply take exhaustive computation time. Furthermore, note that feed-forward neural networks with one hidden layer and lagged process  $\mathbf{Y}_t$  as input acts very similar to an  $\text{AR}(p)$  process if the chosen lags in the input layer are covered in an  $\text{AR}(p)$  process. But as mentioned neural networks are extremely time demanding due to the learning phase that is often based on random selections.

However, we consider three benchmarks from the literature that were applied to electricity prices. First, an  $\text{ARMA}(5,1)$  model with trend as well as annual, weekly and daily cycles as suggested as one of the best models in [Keles et al. \(2012\)](#). Note that their daily cycles vary over the seasons of the year. Moreover, we consider the functional data analysis approach from [Liebl \(2013\)](#). But he is modifying the data as well, as he removes outliers from the data such as holidays and weekends. In fact he is modeling the price by estimating the merit order effect under consideration of the load subtracted by the wind power feed-in. But for our data we noticed that his model gains better results when only the load is used. Furthermore, in his studies two functional principal components were sufficient to model the data well, we got better results by using three principal components. In the original paper the loading coefficients were predicted using a  $\text{SARIMA}$  model with periodicity 5, whereas we use the same model with a periodicity 7 as we are not ignoring the weekends. As a third benchmark from the existent literature we cover the wavelet-ARIMA approach from [Conejo et al. \(2005\)](#). Such a model is often used as a benchmark in electricity spot price forecasting. In our application we use the Daubechies 4 wavelet. For modeling the coefficients of the wavelet decomposition we choose  $\text{ARIMA}(12,1,1)$  processes to capture their autoregressive structure.

As performance measure for our forecasting study we are not considering the MAPE or WMAPE as often done in the literature, because the  $\text{MAPE}_h = \frac{1}{N} \sum_{i=1}^N \left| \frac{Y_{h,i}^{\mathcal{P}} - \hat{Y}_{h,i}^{\mathcal{P}}}{Y_{h,i}^{\mathcal{P}}} \right|$  is obviously pointless for data that can take zeros or even negative values. Instead, we consider the mean absolute forecast error for prediction time  $h$  ( $\text{MAE}_h$ ) as well as the mean of the mean absolute forecast error until forecast horizon  $h$  ( $\text{MMAE}_h$ ). They are given by

$$\text{MAE}_h = \frac{1}{N} \sum_{i=1}^N |\mathbf{Y}_h - \hat{\mathbf{Y}}_{h,i}| \quad \text{and} \quad \text{MMAE}_h = \frac{1}{hN} \sum_{j=1}^h \sum_{i=1}^N |\mathbf{Y}_{t+j} - \hat{\mathbf{Y}}_{j,i}|.$$

The computed  $\text{MAE}_h^{\mathcal{P}}$  and  $\text{MMAE}_h^{\mathcal{P}}$  for the considered models with a prediction horizon of an hour, a day, a week and four weeks are given in Table 8. The evolving of  $\text{MAE}_h^{\mathcal{P}}$  and  $\text{MMAE}_h^{\mathcal{P}}$  with increasing  $h$  is visualised in Figure 10.

It is remarkable that our proposed model can outperform every other one, especially these that are used in literature. Remember that we compare pure out-of-sample methods, that use no information from the prediction time during the estimation. Moreover, we noticed that many models are not able to outperform simple AIC selected  $\text{AR}(p)$  or  $\text{VAR}(p)$  models, so e.g. [Keles et al. \(2012\)](#) and [Liebl \(2013\)](#). Such benchmarks were not considered in most of the literature so far even though there are simple and fast

$h$	$MAE_h^{\mathcal{P}}$				$MMAE_h^{\mathcal{P}}$			
	1	24	168	672	1	24	168	672
pVAR	2.87 (0.28)	7.00 (0.69)	8.88 (0.90)	9.34 (0.81)	2.87 (0.28)	6.07 (0.59)	8.07 (0.79)	9.22 (0.81)
pVAR-TARCH	<b>2.70</b> <b>(0.27)</b>	<b>6.69</b> <b>(0.66)</b>	<b>8.23</b> <b>(0.84)</b>	<b>8.63</b> <b>(0.75)</b>	<b>2.70</b> <b>(0.27)</b>	<b>5.71</b> <b>(0.55)</b>	<b>7.57</b> <b>(0.74)</b>	<b>8.54</b> <b>(0.75)</b>
AR(p)	3.03 (0.30)	6.86 (0.67)	8.66 (0.88)	10.12 (0.87)	3.03 (0.30)	6.09 (0.59)	7.84 (0.77)	9.44 (0.83)
VAR(p)	3.01 (0.30)	6.87 (0.67)	8.66 (0.88)	10.24 (0.89)	3.01 (0.30)	6.09 (0.59)	7.82 (0.77)	9.53 (0.83)
naïve	10.16	10.18	9.84	11.57	10.16	10.38	10.19	11.42
Keles et.al. (2012)	3.21 (0.32)	8.99 (0.88)	9.73 (0.99)	9.53 (0.82)	3.21 (0.32)	7.84 (0.88)	9.39 (0.99)	9.78 (0.82)
Liebl (2013)	10.67 (1.05)	10.48 (1.03)	9.74 (0.99)	10.51 (0.91)	10.67 (1.05)	10.50 (1.01)	12.61 (1.24)	13.83 (1.21)
Conejo et.al. (2005)	7.09 (0.70)	10.81 (1.06)	11.41 (1.16)	13.35 (1.15)	7.09 (0.70)	9.06 (0.87)	11.09 (1.09)	13.20 (1.16)

Table 8:  $MAE_h^{\mathcal{P}}$  and  $MMAE_h^{\mathcal{P}}$  in EUR/MWh of several models (values in parenthesis are % to naïve): pVAR = homoscedastic model (1), pVAR-TARCH = proposed model

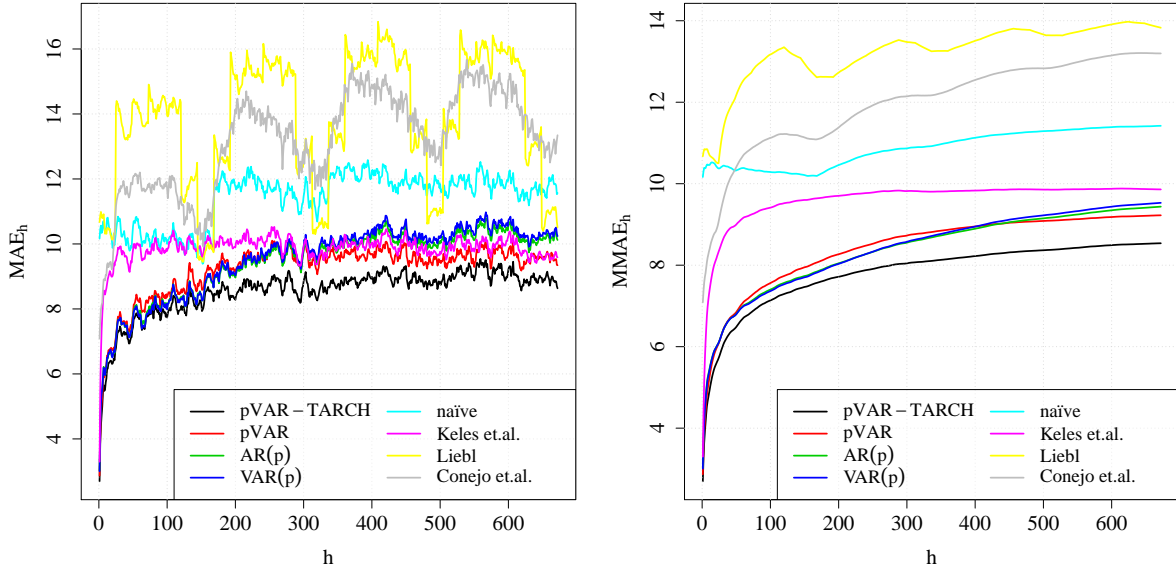


Figure 10:  $MAE_h^{\mathcal{P}}$  and  $MMAE_h^{\mathcal{P}}$  in EUR/MWh of several models for  $h \in \{1, \dots, 672\}$ .

to estimate. The relatively good performance is likely due to the highly chosen model order, so that hundreds of variables are included in the model that can cover the behavior well. This amount of chosen parameters is in the same range as for our reweighted lasso procedure.

Furthermore, we use Monte Carlo methods to compute prediction bands conditioned on the given data. Thus, we can perform a residual based bootstrap on the estimated standardized residuals  $\hat{\mathbf{Z}}_t$  to simulate  $\mathbf{Y}_t$ . So we resample from  $\hat{\mathbf{Z}}_t$  to simulate  $\Sigma_t$  and afterwards  $\mathbf{Y}_t$  to compute prediction intervals.

For illustration purpose we performed a prediction for the sample that is given in Figure 1, which includes a public holiday. So the considered estimation period is about

2 years and the prediction horizon is 192 hours. The corresponding point estimates, such as the estimated conditional mean and 90% resp. 99% prediction intervals are given in figure (11). The prediction bands are computed by evaluating the symmetric conditional value at risk (VaR) levels.

Interestingly the forecasting performs quite well. For example, on Wednesday the 3rd October is an official holiday, so the forecasted values are significantly smaller than for example for the 4th October, even though in this case the forecast still overestimates the real value.

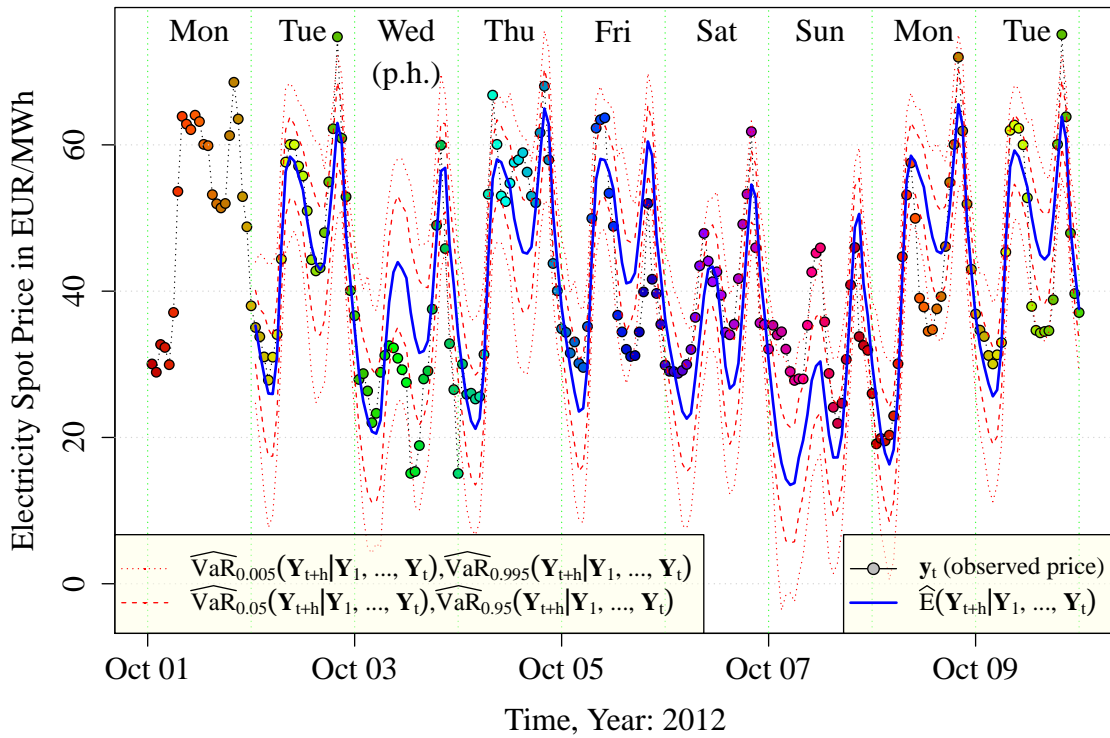


Figure 11: Electricity price prediction with 90% and 99% prediction bands for  $h \in \{1, \dots, 192\}$  using a Monte-Carlo sample size of 1000. The 3rd of October is a public holiday (German unity).

## 7 Summary and Conclusion

Within our paper we modeled the hourly electricity price of the European Energy Exchange for Germany and Austria. A periodic VAR-TARCH approach was proposed in order to capture the specific price movements. This model is able to account for the most difficult challenges in modeling electricity prices. Hence, it considers: the mean-reversion of prices, the seasonality of the data, the possibility of vast price spikes, negative prices, periodicity, time dependent variance, the leverage effect, the impact of renewable energy as well as the impact of electricity load, calendar effects like holidays and time shifts due to daylight saving time.

For the necessary simultaneous modeling of price, load, wind and solar power we used a modern estimation approach, which is efficient and turned out to provide rapid estimations. By fitting our model to the before mentioned data sets we were able to show, that there is a leverage effect for the price and a negative leverage effect for combined

wind and solar power present. Moreover, we provided evidence for the reduction effect of renewable energy on the electricity price. Our study illustrated that the increasing of combined wind and solar power of 1 GWh leads to a decreasing in price of  $0.745 \frac{\text{EUR}}{\text{MWh}}$ .

An extensive forecasting study showed that concerning MAE and MMAE our model outperforms every model which was used as a benchmark, including recently published models in the literature. For the purpose of forecasting no future knowledge was necessary, every forecast was done exclusively with out-of-sample data. Due to its high efficiency and fast computing time our model may be of interest for researchers who wants to use it as a benchmark for their own model as well as for energy companies who needs to forecast the electricity price.<sup>6</sup>

Nevertheless, there is still many research left for future work. For instance, as the markets of some European countries are strongly integrated, measuring the import and export feeds may lead to better estimation results and provide advice for political decisions.

It is also debatable, whether our model can be reasonably applied to other electricity markets. We notice that the characteristics of the the energy portfolio and working day structure of Germany had significant impacts on the price process. But those characteristics are not necessarily true for other markets, therefore introducing a more comprehensive model may be an appropriate future task.

Another direction for future work could also be a more detailed examination of the German energy portfolio, by considering the time series of prices for other energy resources, e.g. coal or gas. Detecting and utilizing the cointegration between those time series may also lead to an improvement for the estimation results.

---

<sup>6</sup>Hence, every reader who is interested in the code, which is written in R, is invited to send a mail to the authors.



## References

- Aggarwal, S. K., Saini, L. M., and Kumar, A. (2009). Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems*, 31(1):13–22.
- Bardet, J.-M., Wintenberger, O., et al. (2009). Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *The Annals of Statistics*, 37(5B):2730–2759.
- Black, F. (1976). Studies in stock price volatility changes. *Proceedings of the Business and Economics Statistics Section. American Statistical Association*, pages 177–181.
- Bordignon, S., Bunn, D. W., Lisi, F., and Nan, F. (2013). Combining day-ahead forecasts for british electricity prices. *Energy Economics*, 35(0):88–103.
- Bosco, B. P., Parisio, L. P., and Pelagatti, M. M. (2007). Deregulated wholesale electricity prices in italy: an empirical analysis. *International advances in economic research*, 13(4):415–432.
- Bowden, N. and Payne, J. E. (2008). Short term forecasting of electricity prices for miso hubs: Evidence from arima-egarch models. *Energy Economics*, 30(6):3186–3197.
- Cancelo, J. R., Espasa, A., and Grafe, R. (2008). Forecasting the electricity load from one day to one week ahead for the spanish system operator. *International Journal of Forecasting*, 24(4):588–602.
- Christensen, T., Hurn, A., and Lindsay, K. (2012). Forecasting spikes in electricity prices. *International Journal of Forecasting*, 28(2):400–411.
- Cifter, A. (2013). Forecasting electricity price volatility with the markov-switching garch model: Evidence from the nordic electric power market. *Electric Power Systems Research*, 102(0):61–67.
- Conejo, A. J., Plazas, M. A., Espinola, R., and Molina, A. B. (2005). Day-ahead electricity price forecasting using the wavelet transform and arima models. *Power Systems, IEEE Transactions on*, 20(2):1035–1042.
- Cruz, A., Muñoz, A., Zamora, J. L., and Espínola, R. (2011). The effect of wind generation and weekday on spanish electricity spot price forecasting. *Electric Power Systems Research*, 81(10):1924–1935.
- Edenhofer, O., Hirth, L., Knopf, B., Pahle, M., Schlömer, S., Schmid, E., and Ueckerdt, F. (2013). On the economics of renewable energy sources. *Energy Economics*, 40(0):12–23.
- Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *The Annals of statistics*, 32(2):407–499.
- Escribano, A., Ignacio Peña, J., and Villaplana, P. (2011). Modelling electricity prices: International evidence. *Oxford bulletin of economics and statistics*, 73(5):622–650.
- Eydeland, A. and Wolyniec, K. (2003). *Energy and power risk management: New developments in modeling, pricing, and hedging*. John Wiley & Sons, New Jersey.

- Fanone, E., Gamba, A., and Prokopczuk, M. (2013). The case of negative day-ahead electricity prices. *Energy Economics*, 35(0):22–34.
- Guthrie, G. and Videbeck, S. (2007). Electricity spot price dynamics: Beyond financial models. *Energy Policy*, 35(11):5614–5621.
- Harvey, A. and Koopman, S. J. (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association*, 88(424):1228–1236.
- Hsu, N.-J., Hung, H.-L., and Chang, Y.-M. (2008). Subset selection for vector autoregressive processes using lasso. *Computational Statistics & Data Analysis*, 52(7):3645–3657.
- Huisman, R. and Kilic, M. (2012). Electricity futures prices: Indirect storability, expectations, and risk premiums. *Energy Economics*, 34(4):892–898.
- Huisman, R., Stradnic, V., and Westgaard, S. (2013). Renewable energy and electricity prices: indirect empirical evidence from hydro power. *IEB Working Paper N. 2013/024*. Available at SSRN.
- Karakatsani, N. V. and Bunn, D. W. (2008). Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, 24(4):764–785.
- Keles, D., Genoese, M., Möst, D., and Fichtner, W. (2012). Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy Economics*, 34(4):1012–1032.
- Keles, D., Genoese, M., Möst, D., Ortlieb, S., and Fichtner, W. (2013). A combined modeling approach for wind power feed-in and electricity spot prices. *Energy Policy*, 59(0):213–225.
- Knittel, C. R. and Roberts, M. R. (2005). An empirical examination of restructured electricity prices. *Energy Economics*, 27(5):791–817.
- Koopman, S. J., Ooms, M., and Carnero, M. A. (2007). Periodic seasonal reg-arfima-garch models for daily electricity spot prices. *Journal of the American Statistical Association*, 102(477):16–27.
- Kristiansen, T. (2012). Forecasting nord pool day-ahead prices with an autoregressive model. *Energy Policy*, 49(0):328–332.
- Lawson, C. L. and Hanson, R. J. (1995). *Solving least squares problems*. SIAM.
- Liebl, D. (2013). Modeling and forecasting electricity spot prices: A functional data perspective. *The Annals of Applied Statistics*, 7(3):1562–1592.
- Liu, H. and Shi, J. (2013). Applying arma-garch approaches to forecasting short-term electricity prices. *Energy Economics*, 37(0):152–166.
- Mak, T., Wong, H., and Li, W. (1997). Estimation of nonlinear time series with conditional heteroscedastic variances by iteratively weighted least squares. *Computational statistics & data analysis*, 24(2):169–178.

- Mbamalu, G. and El-Hawary, M. (1993). Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation. *Power Systems, IEEE Transactions on*, 8(1):343–348.
- Rabemananjara, R. and Zakoïan, J. M. (1993). Threshold arch models and asymmetries in volatility. *Journal of Applied Econometrics*, 8(1):31–49.
- Ren, Y. and Zhang, X. (2010). Subset selection for vector autoregressive processes via adaptive lasso. *Statistics & probability letters*, 80(23):1705–1712.
- Sensfuß, F., Ragwitz, M., and Genoese, M. (2008). The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in germany. *Energy policy*, 36(8):3086–3094.
- Svrien, N. and Eric, Y. Y. L. (2012). Shrinkage and lasso strategies in high dimensional heteroscedastic models. Technical report, Report, University of Windsor, Department of Mathematics and Statistics.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58(1):267–288.
- Wagener, J. and Dette, H. (2013). The adaptive lasso in high-dimensional sparse heteroscedastic models. *Mathematical Methods of Statistics*, 22(2):137–154.
- Wang, H., Li, G., and Tsai, C.-L. (2007). Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(1):63–78.
- Weron, R. (2006). *Modeling and forecasting electricity loads and prices: A statistical approach*. John Wiley & Sons, Chichester.
- Woo, C.-K., Horowitz, I., Moore, J., and Pacheco, A. (2011). The impact of wind generation on the electricity spot-market price level and variance: The texas experience. *Energy Policy*, 39(7):3939–3944.
- Würzburg, K., Labandeira, X., and Linares, P. (2013). Renewable generation and electricity prices: Taking stock and new evidence for germany and austria. *Energy Economics*, 40(0):159–171.
- Yan, X. and Chowdhury, N. A. (2013). Mid-term electricity market clearing price forecasting: A hybrid lssvm and armax approach. *International Journal of Electrical Power & Energy Systems*, 53(0):20–26.
- Ziel, F. and Schmid, W. (2013). Quasi-maximum likelihood estimation of periodic autoregressive, conditionally heteroscedastic time series. Technical report, Discussion Paper, European University Viadrina, Department of Business Administration and Economics.